

EARS: New Directions

Julia Hirschberg
Columbia University
14 November 2003

Rich Transcription

- Goal: improve usability of SST
 - Usability = Readability? Browsability? Searchability?
- RT is
 - Speaker diarization
 - SU detection
 - Disfluency identification

Richer Transcription

- Speaker diarization [Speaker characterization
 - Speaker identification
 - Speaker characteristics (gender, age, state, type)
- SU detection [Structure detection
 - Type labeling
 - Paratone detection, section labeling
 - Capitalization (e.g. of proper names)
- Disfluency identification [Disfluency handling
 - Delete reparanda or ... identify but ...

RT needs for BN vs CTS

- For readability or browsability or searchability?
 - Speaker characterization
 - CTS: ID, gender, age, language
 - BN: ID, type
 - Structural identification
 - CTS: SU, paratone, speaker turns
 - BN: SU, paratone, section labels, capitalization
 - Disfluency handling
 - CTS: more
 - BN: fewer

Evaluation

- Couple error analysis with usability:
 - Which errors hurt usability most?
- Task-based evaluation of usability
 - What tasks will users want to perform?
- Discourse structure segmentation may provide some ideas for SU detection scoring
 - 2-level labeling and evaluation measures